Hacker Highschool
SECURITY AWARENESS FOR TEENS

LEÇON 8
ANALYSE LÉGALE DANS LE MONDE NUMÉRIQUE

Hacker Highschool
SECURITY AWARENESS FOR TEENS

ISECOM
INSTITUTE FOR SECURITY AND OPEN METHODOLOGIES

“License for Use” Information

The following lessons and workbooks are open and publicly available under the following terms and conditions of ISECOM:

All works in the Hacker Highschool project are provided for non-commercial use with elementary school students, junior high school students, and high school students whether in a public institution, private institution, or a part of home-schooling. These materials may not be reproduced for sale in any form. The provision of any class, course, training, or camp with these materials for which a fee is charged is expressly forbidden without a license including college classes, university classes, trade-school classes, summer or computer camps, and similar. To purchase a license, visit the LICENSE section of the Hacker Highschool web page at www.hackerhighschool.org/license.

The HHS Project is a learning tool and as with any learning tool, the instruction is the influence of the instructor and not the tool. ISECOM cannot accept responsibility for how any information herein is applied or abused.

The HHS Project is an open community effort and if you find value in this project, we do ask you support us through the purchase of a license, a donation, or sponsorship.

All works copyright ISECOM, 2004.

Conditions d’utilisation de ce support

Ces leçons et supports sont gratuits et disponibles pour le public sous les conditions suivantes d’ISECOM:

Tous les travaux menés dans le cadre du "Hacker HighSchool" sont disponibles à usage non commercial auprès d’élèves du collège, du lycée, dans le cadre d’écoles publiques ou privées, ou encore lors de scolarisations à domicile. Ces supports ne peuvent être reproduits en vue d’un usage commercial. Il est expressément interdit d’utiliser ces supports dans le cadre de cours, leçons et/ou stages payants, à moins d’obtenir une licence pour cela (dans ce cas, veuillez aller sur *www.hackerhighschool.org/license*).

Le projet HSS (Hacker HighSchool) est un outil de travail et d’apprentissage, et en tant que tel, son utilisation relève de la personne qui l’utilise, et non de l’outil lui-même. ISECOM ne peut être mis en cause si cet outil est utilisé à mauvais escient ou de manière illégale.

Le projet HSS est aussi le fruit de l’effort de toute une communautés, et si vous trouvez ce projet intéressant, nous vous serions plus que reconnaissants de votre aide, soit par l’achat d’une licence, soit par un don, soit encore par un quelconque parrainage.

Table des matières

Personnes ayant contribué à ce projet

Simon Biles, Computer Security Online Ltd.

Pete Herzog, ISECOM

Chuck Truett, ISECOM

Marta Barceló, ISECOM

Kim Truett, ISECOM

Traduction

Bénoni Martin

ISECOM
INSTITUTE FOR SECURITY AND OPEN METHODOLOGIES

/// computer security online

8.0 Introduction

L'analyse légale consiste à utiliser une technique d'investigation méthodique dans l'objectif de reconstruire une suite d'évènements. Le concept de médecin légiste est familier pour la plupart grâce à la télévision et aux films, en particulier pour les investigations des scènes de crime. La science légiste était longtemps – et l'est toujours dans la réalité – associée avec la médecine légiste, c'est-à-dire déterminer comment les gens sont morts. La première trace décrivant une analyse légale portait sur ce sujet. En 1248, un livre chinois Hsi Duan Yu (littéralement *le lessivage des des erreurs*) a été publié. Ce livre décrit comment distinguer si une personne est morte de noyade ou de strangulation[1].

L'analyse légale dans le monde numérique est un petit peu plus confus et moins connu. C'est l'aptitude à reconstruire ce qui s'est passé sur un appareil numérique. Auparavant, cela était restreint aux ordinateurs mais cela comprend actuellement tous les appareils numériques comme les téléphones portables, les appareils photos numériques, et même les appareils GPS[2]. Ces techniques ont été utilisés pour retrouver des meurtriers, kidnappeurs, fraudeurs, des parains de la Mafia et encore d'autres personnes malveillantes.

Dans cette leçon, nous allons développer deux aspects de l'analyse légale (dans le domaine des ordinateurs, nous n'aborderons pas ici les téléphones portables).

1. Ce qui se trouve sur les ordinateurs personnels
Cela couvre
- la récupération des fichiers effacés
- la cryptanalyse élémentaire
- la recherche de certains types de fichiers
- la recherche de certains motifs particuliers
- l'examen de certaines zones intéressantes de l'ordinateur

2. Ce qu'un utilisateur a fait à distance sur l'ordinateur d'une autre personne
Cela couvre
- la lecture des fichiers journaux
- la reconstruction des actions réalisées
- le traçage de la source

Cette leçon va se focaliser sur les outils disponibles sous Linux. Il existe aussi des outils disponibles sous Windows, ainsi que des logiciels ou matériels dédiés à l'analyse légale, mais grâce à la capacité de Linux de monter et de lire un nombre élevé de systèmes d'exploitation et de fichiers, c'est l'environnement idéal pour la plupart des opérations d'analyse légale.

1 De tout évidence, les marques autour du cou, ainsi que le niveau de pénétration de l'eau dans les poumons sont utilisés.
2 Global Positioning System – système de géolocalisation par satellite.

8.1 Principes de l'analyse légale

8.1.0 Introduction

Que vous examiniez un ordinateur ou un corps, certains principes de bases sont nécessaires. Cette section est un résumé de ces principes.

8.1.1 Eviter la contamination

A la télévision, vous voyez les médecins légistes en blouse blanche avec des gants, prenant les preuves à l'aide de pinces à épiller et les conservant dans des sacs en plastique scellés. Cela est fait pour éviter la «contamination». Cela arrive lorsqu'une preuve est par exemple souillée par des empreintes digitales lorsqu'un couteau est ramassé (Voir le film *le fugitif* … et à quelles difficultés cela le conduit!)

8.1.2 Agir méthodiquement

Quoique vous faîtes, lorsque (si?) dans l'éventualité où vous devriez aller au tribunal, vous allez devoir justifier toutes les actions réalisées. Si vous agissez d'une manière méthodique et scientifique, en prenant des notes sur ce que vous faites et comment vous le faites, cette justification devient plus facile. Cela permet aussi pour quelqu'un d'autre qui suit votre démarche, de vérifier que vous n'avez pas fait d'erreur qui peut mettre en doute votre preuve.

8.1.3 Le chemin de la preuve

Vous devez entretenir ce qu'on appelle le «chemin de la preuve». Cela signifie qu'en tout point depuis la saisie de la preuve jusqu'à sa présentation finale au tribunal, vous devez pouvoir dire qui a accédé à cette preuve et où elle est conservée. Cela permet d'éviter que quelqu'un manipule ou falsifie la preuve de quelque manière que ce soit.

8.1.4 Conclusion

Gardez ces principes à l'esprit, et même si vous n'allez pas jusqu'au tribunal, vous pourrez maximiser vos aptitudes d'analyste «légiste».

8.2 Analyse légale d'un ordinateur autonome

8.2.0 Introduction

Cette section aborde l'analyse légale d'un seul ordinateur. Pour plus de compréhension, nous l'appellerons «l'analyse légale d'un ordinateur autonome». C'est probablement la partie la plus commune de l'analyse légale dans le monde numérique – son rôle principal est de trouver ce qui a été fait en utilisant un ordinateur particulier. L'analyste «légiste» peut rechercher la preuve d'une fraude, comme des feuilles de calcul financières, la preuve d'une communication avec quelqu'un, des messages électroniques ou un carnet d'adresses, ou une preuve d'une nature particulière, comme la présence d'images pornographiques.

8.2.1 Principes de base du disque dur et des media de stockage

Un ordinateur regroupe plusieurs composants. Ce sont le processeur, la mémoire, la carte graphique, les lecteurs CD et bien d'autres. L'un des composant les plus importants est le disque dur. C'est l'endroit dans lequel la majorité des informations nécessaires au fonctionnement de l'ordinateur sont stockées. Le Système d'Exploitation (SE ou OS pour *Operating System* en anglais) comme Windows ou Linux y est installé, ainsi que les applications comme le traitement de texte ou des jeux. C'est aussi l'endroit où de nombreuses données sont stockés de manière volontaire, comme suite à l'action de sauvegarder un fichier, ou de manière non intentionnelle, par l'utilisation de fichiers temporaires et de cache. Cela permet à l'analyste «légiste» de reconstruire les actions que l'utilisateur a menées sur une ordinateur, quels fichiers ont été accédés, et beaucoup plus.
Il existe plusieurs niveaux auquels on peut examiner un disque dur. Pour les besoins de l'exercice, nous allons uniquement regarder le niveau du système de fichiers. Il est cependant bon de noter que les professionnels sont capables de regarder de manière très précise un disque dur afin de déterminer son contenu – même suite à plusieurs réécritures successibes.
Le système de fichiers est l'application à un ordinateur des classeurs. Il contient des tiroirs (partitions), des dossiers (répertoires) et des feuilles de papiers (fichiers). Les fichiers et répertoires peuvent être cachés, cependant c'est une fonction qui peut être contournée très facilement.
Les exercices suivants vous donneront une meilleure connaissance des principes du stockage de disque.

Exercices:
Pour chacun des termes suivants concernant les media de stockage, recherchez comment ils fonctionnent. La première étape de l'analyse légale consiste à comprendre comment les équipements fonctionnent en situation normale.
1. Disque magnétique/dur/physique: C'est l'endroit où votre ordinateur stocke les fichiers. Expliquez comment le magnétisme est utilisé dans un disque dur.
2. Tracks: Qu'appelle-t-on les tracks d'un disque dur?
3. Secteurs: C'est un espace de taille fixe dans lequel les données sont positionnées. Expliquez comment.
4. Cluster/unité d'allocation: Expliquez pourquoi lors de l'écriture d'un fichier sur le disque dur, l'espace alloué peut être supérieur à ce qui est nécessaire. Qu'est-ce qu'il advient de cet espace vide? Rechercher le terme «*file slack*» peut vous aider.

5. Espace disponible/non alloué: C'est ce que vous laissez lorsque des fichiers sont effacés. Est-ce que ces fichiers ont réellement disparus? Expliquez comment un fichier est effacé sur l'ordinateur. Rechercher des outils d'« effacement sécurisé» peut vous aider. Savoir comment vous devez supprimer de manière sécurisée un fichier pour qu'il soit vraiment supprimé est un bon moyen d'apprendre pourquoi ces outils sont nécessaires.
6. Hash et Hash MD5 : Expliquez ce qu'est un hash et pourquoi est-il utilisé?
7. BIOS: Cela signifie «Basic Input/Output System» (système basique d'entrée/sortie). Qu'est-ce que c'est et où est ce que c'est stocké dans l'ordinateur?
8. Secteur d'amorçage: Il travaille avec les tables de partitions pour que votre PC trouve le système d'exploitation à lancer. Il existe de nombreux outils sur les partitions, le standard étant fdisk. La connaissance du fonctionnement de ces outils est la première indication pour comprendre comment fonctionnent les partitions et le secteur d'amorçage.
9. Contrôle de Redondance Cyclique (CRC): Lorsque vous obtenez une «erreur de lecture» de votre disque dur, cela signifie que le contrôle CRC a échoué. Recherchez ce qu'est le contrôle CRC et ce qu'il fait.
10. Signature de fichiers: Souvent les fichiers ont une signature courte de 6 octets au début indiquant le type de fichier. Ouvrir un fichier dans un éditeur de texte est la manière la plus simple de le voir. Ouvrez 3 fichiers de chacun des types suivantsdans un éditeur de texte : . jpg, .gif, .exe, .mp3. Quel est le premier mot en tête de chaque fichier?
11. RAM (Random-Access Memory – mémoire vive): C'est aussi connu sous le nom «mémoire» et c'est une localisation temporaire pour écrire et lire de l'information. C'est beaucoup plus rapide que l'écriture sur le disque dur. Le contenu est perdu lors de l'arrêt de l'ordinateur. Expliquez comment la RAM fonctionne. Sachant que votre ordinateur peut avoir de 64 à 512Mb de RAM, recherchez de l'information sur un ordinateur qui dispose de plus de RAM.
12. Actuellement le plus grand disque RAM (disque dur ultra-rapide émulé en RAM) est 2,5Tb (Tera octets). Cela est combien de fois plus grand que votre ordinateur?

8.2.2 Chiffrement, déchiffrement et formats de fichiers

De nombreux fichiers rencontrés peuvent ne pas être immédiatement lisibles. De nombreux programmes ont leur propre format de fichiers propriétaire, tandis que d'autres utilisent des formats standards – par exemple les formats standards d'images comme gif, jpeg, etc. Linux fournit un excellent utilitaire pour vous aider à déterminer le format d'un fichier. C'est l'outil file.

Option de ligne de commande	Résultat
-k	Don't stop at the first match, keep going.
-L	Follow symbolic links
-z	Attempt to look inside compressed files.

Un exemple d'utilisation de la commande file est montré ici:

```
[simon@frodo file_example]$ ls
arp.c                   nwrap.pl
isestorm_DivX.avi       oprp_may11_2004.txt
krb5-1.3.3              VisioEval.exe
```

```
krb5-1.3.3.tar          Windows2003.vmx
krb5-1.3.3.tar.gz.asc
[simon@frodo file_example]$ file *
arp.c:                          ASCII C program text
isestorm_DivX.avi:              RIFF (little-endian) data, AVI
krb5-1.3.3:                     directory
krb5-1.3.3.tar:                 POSIX tar archive
krb5-1.3.3.tar.gz.asc:          PGP armored data
nwrap.pl:                       Paul Falstad's zsh script text
executable
oprp_may11_2004.txt:            ASCII English text, with very long
lines, with CRLF line terminators
VisioEval.exe:                  MS-DOS executable (EXE), OS/2 or MS
Windows
Windows2003.vmx:                a /usr/bin/vmware script text
executable
[simon@frodo file_example]$
```

A partir de cela, vous pouvez faire des tentatives de lectures sur certains types de fichiers. Il y a de nombreux outils de conversion de types de fichiers sous Linux et encore plus sur Internet, ainsi que des visualiseurs de fichiers pour différents formats. Parfois plusieurs étapes sont nécessaires pour pouvoir lire des données, essayez de progresser pas à pas!

Il arrive de tomber sur des fichiers chiffrés ou protégés par mot de passe. Cela représente des difficultés variants d'un chiffrement facilement cassable, à ce qui peut poser des problèmes même à des agences gouvernementales comme la NSA. Il y a là encore de nombreux outils disponibles sur Internet que vous pouvez utilisez pour essayer de déchiffrer un fichier. Il est aussi utile d'examiner l'ordinateur dans son ensemble. Il est difficile de se souvenir des mots de passe, ils peuvent être écrits quelque part. Les choix classiques pour les mots de passe concernent: les animaux familiers, les proches, les dates (mariage, naissance), numéros de téléphone, plaques minéralogiques, et d'autres combinaisons (123456, abcdef, azerty, etc). Les gens généralement refusent d'utiliser plus d'un ou deux mots de passe en tout, donc si vous découvrez un mot de passe pour un fichier ou une application, vous pouvez l'utiliser sur les autres. C'est probablement le même.

Exercices:
Vous apprendrez le cassage de mots de passe. Il est légal de casser ses propres mots de passe en cas de perte, ce n'est peut-être pas légal dans certains pays de rechercher comment quelque chose est chiffré, afin de protéger les autres équipements.
Les films DVD sont chiffrés afin de les protéger du vol et de la revente illégale. C'est un excellent exemple d'utilisation du chiffrement mais c'est illégal de rechercher comment le chiffrement est utilisé. Cela nous mène au premier exercice:
1. Qu'est-ce que «DeCSS» et comment est-ce lié au chiffrement DVD? Recherchez «decss» pour en savoir plus.
2. Sachant qu'un fichier est protégé par mot de passe, il est nécessaire de trouver comment ouvrir ce fichier. C'est ce qu'on appelle «casser» un mot de passe. Recherchez des informations sur le cassage de différents types de mots de passe. Vous pouvez recherchez

«cracking XYZ passwords» (cassage des mots de passe XYZ) où XYZ est le type de mot de passe recherché. Faîtes le avec les types de mots de passe suivants:
a. MD5
b. Adobe PDF
c. Excel
3. Si la méthode de chiffrement est trop forte pour être cassée, il peut être nécessaire de faire une attaque par dictionnaire (connue aussi sous le nom attaque par force brute). Trouvez en quoi consiste une attaque par dictionnaire.

8.2.3 Trouver une aiguille dans une botte de foin

Les logiciels commerciaux d'analyse légale incluent de puissants outils de recherche permettant de trouver des combinaisons et permutations de facteurs. Sans ces outils commerciaux, vous devez être un peu plus malin. Linux permet de construire des outils similaires à l'aide des utilitaires standards. Le paragraphe suivant décrit l'utilisation de find, grep et strings, puis décrit l'utilisation de pipe pour les combiner.

8.2.3.1 find

find est utilisé pour localiser les fichiers répondant à certains critères au sein du système d'exploitation. Il ne permet pas de rechercher à l'intérieur des fichiers. Il peut y avoir beaucoup de permutations d'expressions qui peuvent être combinés pour recherche un fichier.
Exercice:
1. Lisez le manuel de find. Complétez le résultat de chaque expression dans le tableau suivant. (Indice: Lorsqu'on nombre est donné comme argument, il peut être spécifié comme suit: +n – plus grand que n; -n – plus petit que n; n – exactement n).

Expression	Résultat
-amin n	Fichiers étant accédés il y a n minutes
-anewer	
-atime	
-cnewer	
-iname	
-inum	
-name	
-regex	
-size	
-type	
-user	

8.2.3.2 grep

grep est un outil très puissant. Il est utilisé pour trouver certaines lignes au sein d'un fichier. Cela permet de rapidement trouver les fichiers particuliers au sein d'un répertoire ou système de fichiers. Il permet la recherche à l'aide des expressions régulières. On peut utiliser des motifs qui correspondent à ce qu'on cherche. Par exemple: trouver toutes les chaînes de caractères qui commencent par «s» et se terminent par «t» afin de finir des mots-croisés.

Exercices:
1. Lisez le manuel de grep
2. Recherchez des expressions régulières pour grep sur Internet. Essayez de construire une expression régulière qui cherche tous les mots de quatre lettres qui contiennent un «a».

8.2.3.3 strings

strings est un autre utilitaire utile. Il permet de chercher dans un fichier toutes les chaînes de caractères lisibles par une personne. Cela peut retourner un nombre important d'informations sur un fichier spécifique, souvent sur l'application qui a créé le fichier, les auteurs, la date de création et bien d'autres.

Exercice:
1. Lisez le manuel de strings

8.2.3.4 awk

awk est un langage de programmation destiné à la manipulation des chaînes de caractères. Il est utilisé pour extraire l'information d'une commande et l'envoyer dans une autre. Par exemple, pour afficher uniquement les programmes en cours d'exécution depuis la commande ps, vous pouvez l'utiliser comme suit:

Exercice:
1. Lisez le manuel de awk

8.2.3.5 pipe

Tous les outils ci-dessous sont combinés facilement grâce à la commande UNIX «pipe». Elle est utilisée sous le symbole «|». Elle permet de prendre le résultat d'une commande et de le mettre en entrée d'une autre commande. Pour trouver tous les fichiers mpg du répertoire courant, vous pouvez l'utiliser comme suit:

Exercices:
1. En utilisant la pipe, les commandes ls et grep, listez tous les fichiers dans le répertoire courant qui ont été créés ce mois.
2. En utilisant les commandes ps et awk, listez les noms des process en cours d'exécution.

8.2.4 Utiliser d'autres sources

Il y a bien d'autres moyens de découvrir comment un ordinateur est utilisé. Presque chaque application garde une trace des fichiers qu'elle traite. Cela inclus les fichiers temporaires, la liste des fichiers précédemment accédés, ou l'historique d'un navigateur.

Exercices:

1. Qu'est ce que le cache d'un navigateur? Trouvez où le navigateur stocke son cache.
2. Que sont les cookiesd'un navigateur ? Trouvez où le navigateur les stocke.
3. Trouvez des informations au sujet des cookies du navigateur web. Quel type de cookies sont présents et quels types d'information y sont stockés?
4. Votre ordinateur utilise des répertoires temporaires dans lesquels il écrit par défaut des fichiers pour l'utilisateur. C'est souvent connu sous le nom «Application Data». Trouver les répertoires temporaires sur votre ordinateur. Ils sont souvent appelés tmp ou temp mais ce ne sont pas les seuls. Essayez de trouver les fichiers écrit à la date d'aujourd'hui est un bon moyen de découvrir les fichiers temporaires. Est-ce que ces fichiers disparaissent lorsque vous redémarrez l'ordinateur?

8.3 Analyse légale réseau

8.3.0 Introduction

L'analyse légale réseau est utilisée pour découvrir où un ordinateur est situé et pour prouver qu'un fichier particulier a été envoyé depuis un ordinateur particulier. L'analyse légale réseau peut être très compliquée, cependant nous allons couvrir des principes qui peuvent être appliquées de manière générale.

8.3.1 Journaux du coupe feu

Qui se connecte sur mon ordinateur? Le coupe feu est l'utilitaire qui peut interdire les connexions entre deux points d'un réseau. Plusieurs types de coupe feu existent. Quelque soit le type du coupe feu, les journaux fournissent des informations détaillées. Rien qu'en lisant les journaux, vous pouvez repérer des attaques types sur votre coupe feu.

Exercices:
1. Visitez http://www.dshield.org. Ce site analyse les journaux de coupe feu du monde entier pour découvrir les tentatives d'attaques réseau. Cela aide les professionnels de la sécurité pour vérifier la protection du réseau par rapport à des vulnérabilités avant qu'elles ne surviennent. Lisez le site et expliquez comment le graphique mondial est obtenu et sa signification.
2. Sur le même site, lisez la section «Fight back» et les messages de réponse reçus. Expliquez en l'objectif.

8.3.2 En-tête de messages

Les messages gardent la trace de toutes les machines traversées pour arriver jusqu'à votre boîte aux lettres. Cela est conservé dans l'en-tête du message. D'autres informations y sont aussi disponibles. Suivant le client de messagerie, il est plus ou moins facile d'accéder à ces informations. Il faut garder à l'esprit que l'information est indiquée à l'envers. En haut de la liste vous avez votre ordinateur. Ensuite, chaque ligne jusqu'à la dernière indique l'ordinateur ou le réseau par lequel le message a été transmis.

Exercices:
1. Une ressource importante sur la lutte contre le SPAM est http://www.spamspade.org. Visitez le site dans la section «The Library». En utilisant cette section, vous serez capables de comprendre la signification des en-têtes. Vous pouvez aussi vous renseigner sur les abus et les messages forgés. Expliquez les différents cas d'utilisation malicieuse de la messagerie électronique.
2. Trouvez comment accéder aux en-têtes de vos messages. Y a-t-il des champs particuliers que vous ne connaissez pas? Recherchez leur signification. Vous devez pouvoir expliquer chaque champ des en-têtes.

Pour en savoir plus

The following links are in English.
http://www.honeynet.org/papers/forensics/
http://www.honeynet.org/misc/chall.html - Some forensic Exercises.
http://www.porcupine.org/forensics/ - The classics
http://www.computerforensics.net/
http://www.guidancesoftware.com/corporate/whitepapers/index.shtm#EFE
http://www.forensicfocus.com/
http://www.securityfocus.com/infocus/1679
http://www.linuxsecurity.com/feature_stories/feature_story-139.html
http://www.linuxsecurity.com/feature_stories/feature_story-140.html
http://www.securityfocus.com/incidents
http://staff.washington.edu/dittrich/talks/blackhat/blackhat/forensics.html
http://www.openforensics.org/
http://fire.dmzs.com/
http://www.sleuthkit.org/
http://www.fbi.gov/hq/lab/fsc/backissu/oct2000/computer.htm